MATHEMATICS

HYPERGEOMETRIC FUNCTIONS. I

BY

A. H. M. LEVELT

(Communicated by Prof. N. G. DE BRUIJN at the meeting of November 26, 1960)

PREFACE

The hypergeometric series is defined by

$$(1)\qquad \mathsf{F}\left(\begin{matrix} d_1, \ldots, d_n \\ e_1, \ldots, e_{n-1} \end{matrix}\middle| z\right) = \sum_{j=0}^{\infty} \frac{(d_1)_j \ldots (d_n)_j}{j!\,(e_1)_j \ldots (e_{n-1})_j}\, z^j,$$

where $(d)_0=1$ and $(d)_j=d(d+1)\ldots(d+j-1)$ if $j\geqslant 1$, and $d_1, \ldots, d_n, z$ are complex numbers, $e_1, \ldots, e_{n-1}$ are complex numbers $\neq 0, -1, -2, \ldots$. When $n>2$ the series is often called a generalized hypergeometric series. The hypergeometric function is the complete analytic continuation of the function defined inside the unit circle by the series (1). The special case $n=2$ (then $\mathsf{F}\left(\begin{matrix} a, b \\ c \end{matrix}\middle| z\right)$ is usually written as $F(a, b; c; z)$) has been extensively studied by Euler, Gauss, Kummer and Riemann. We do not intend to expound the work of these authors (For this we refer to F. KLEIN [1]) [1]). We shall just indicate the methods of KUMMER [2] and RIEMANN [3]. Kummer's paper contains the first systematic investigation of the second-order differential equation satisfied by $F(a, b; c; z)$. He found many new results, such as the well-known 24 series expansions and the quadratic transformation formulas. However, the calculations due to the transformations of the differential equation are rather lengthy, and it is not quite clear why some transformations lead to interesting relations whereas others do not.

Riemann considered a two-dimensional space V (over the complex number field) of functions which are holomorphic on the universal covering surface W of Z'. Here Z' denotes the complex number sphere from which three points a, b and c are removed. He showed that V is the solution space of a hypergeometric differential equation, if the following conditions are satisfied:

(i) V contains functions a_1, a_2; b_1, b_2 and c_1, c_2, which can be written as $(z-a)^{\alpha_1} f_1(z)$, $(z-a)^{\alpha_2} f_2(z)$; $(z-b)^{\beta_1} g_1(z)$, $(z-b)^{\beta_2} g_2(z)$ and $(z-c)^{\gamma_1} h_1(z)$, $(z-c)^{\gamma_2} h_2(z)$ in neighbourhoods of a, b and c, respectively.

[1]) Numbers in square brackets refer to the references, which are given at the end of the preface and at the end of each chapter.

(ii) $f_1, f_2, g_1, g_2, h_1, h_2$ are holomorphic and non-vanishing at the points a, b and c, respectively.

(iii) $\alpha_1-\alpha_2$, $\beta_1-\beta_2$ and $\gamma_1-\gamma_2$ are not integers, and $\alpha_1+\alpha_2+\beta_1+\beta_2+\gamma_1+\gamma_2=1$. (We shall call α_1, α_2; β_1, β_2 and γ_1, γ_2 the exponents at a, b and c, respectively).

This approach of Riemann's appeared to be very fruitful. He was able to deduce most of Kummer's results hardly without calculations. He also determined the linear transformations A, B and C of V induced by the analytic continuation of the functions of V around the points a, b and c, respectively (The group generated by A, B and C is a representation of the fundamental group of Z', and is called the monodromy group). The linear relations between the three bases $\{a_1, a_2\}$, $\{b_1, b_2\}$ and $\{c_1, c_2\}$ of V then follow at once.

For the first time generalized hypergeometric functions were studied systematically by J. THOMAE [4]. He regarded these functions as solutions of an n-th order differential equation of Fuchsian type having singularities at the points $0, 1, \infty$.

Following the method of Riemann's paper E. GOURSAT [5] defined generalized hypergeometric functions not as solutions of a certain differential equation, but by properties in the large. However, he did not try to determine the monodromy group.

At the end of the last century integral representation of hypergeometric functions were found; they have been extensively used ever since (see e.g. [6] also for references).

In more recent times hypergeometric functions were almost exclusively considered from the point of view of differential equations. Many kinds of special solutions have been obtained, and the relations between them have been deduced analytically. Riemann's program, partly carried out by Goursat, was not completed until now. This slow progress was possibly caused by the difficulties which arise in the case where some differences of exponents are equal to integers. This problem can be solved if one succeeds to characterize the solution space of an analytic differential equation in a neighbourhood of a regular (or weakly) singular point. We shall do this in chapter II, in which a new theory of regular singular points will be given. It is a general theory which is by no means restricted to the hypergeometric differential equation.

The monodromy group of the generalized hypergeometric function can be found in a purely algebraic way, after a certain simple algebraic problem has been solved. Chapter I is devoted to the solution of this problem. Hypergeometric functions will be defined in chapter III under the least possible restrictions on the exponents.

The definition is analogous to Riemann's, which was described above. In the same chapter we shall show that the hypergeometric function satisfies a differential equation, and we shall compute the monodromy group, using the results of chapter I.

Finally, in chapter IV we shall describe some examples of applications.

REFERENCES

1. KLEIN, F., Vorlesungen über die hypergeometrische Funktion. Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen. Band XXXIX. Springer, 1933.
2. KUMMER, E., Über die hypergeometrische Reihe $1+\frac{\alpha\cdot\beta}{1\cdot\gamma}x+\ldots$. J. reine angew. Math. **15** (1836), 39–83 und 127–172.
3. RIEMANN, B., Beiträge zur Theorie der durch die Gauss'sche Reihe $F(\alpha, \beta, \gamma, x)$ darstellbaren Functionen. Gesammelte mathematische Werke. Herausgegeben von H. Weber. 67–87. Dover Publications, Inc. New York.
4. THOMAE, J., Über die höheren hypergeometrischen Reihen, insbesondere über die Reihe:
$$1+\frac{a_0\, a_1\, a_2}{1\cdot b_1\, b_2}\,x+\frac{a_0(a_0+1)\,a_1(a_1+1)\,a_2(a_2+1)}{1\cdot 2\cdot b_1(b_1+1)\,b_2(b_2+1)}\,x^2\ldots$$
Math. Ann. 2 (1870), 427–444.
5. GOURSAT, E., Mémoire sur les fonctions hypergéometriques d'ordre supérieur. Ann. sci. école norm. sup. (2), **12** (1883), 261–286, 395–430.
6. NØRLUND, N. E., Hypergeometric Functions. Acta Math. **94** (1955), 289–349.

A PRELIMINARY ALGEBRAIC PROBLEM

I. 1. *Introduction*

Since J. THOMAE's paper [1] of 1870, which initiated the theory of generalized hypergeometric functions, many papers on this subject have appeared. A systematic exposition of the theory and many reference can be found in [2]. However, it has never been noticed that the properties of these functions are dominated by a simple algebraic structure, embodied in a problem which will be posed and solved in the sequel.

In this first chapter we deal with this algebraic problem. In chapter III we shall give the analytic foundations of the theory, and only then we shall be able to describe the rôle of the algebraic problem in the general theory of hypergeometric functions. The combination of algebraic and analytic theories will lead to most of the well known formulas for hypergeometric functions both ordinary and generalized.

I. 2. *The algebraic problem*

Let n be an integer $\geqslant 2$ (if $n=1$, the exposition which follows, although making sense, becomes trivial). Let V be an n-dimensional vector space over a (commutative) field L. Let

$$(1.1)\qquad P(x)=x^n+p_1x^{n-1}+\ldots+p_n,\; Q(x)=x^n+q_1x^{n-1}+\ldots+q_n$$

be polynomials with coefficients in L. The first version of our problem is:

Problem I

a. Do there exist linear transformations A and B of V satisfying the following conditions:
   (1) $P(x)$ and $Q(x)$ (see (1.1)) are the characteristic polynomials of A and B, respectively.
   (2) rank $(A-B) \leqslant 1$.

b. On what conditions and in what sense has problem Ia a unique solution?

The answer to problem Ia is affirmative. I owe the following very simple proof to Prof. N. G. DE BRUIJN. Let A_1 and B_1 be the companion matrices ([3], p. 318) of $P(x)$ and $Q(x)$, respectively:

$$(1.2)\quad A_1 = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -p_n \\ 1 & \cdot & & & & & \cdot & \cdot \\ 0 & \cdot & \cdot & & & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot & 0 & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 & -p_1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -q_n \\ 1 & \cdot & & & & & \cdot & \cdot \\ 0 & \cdot & \cdot & & & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot & 0 & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 & -q_1 \end{pmatrix}.$$

Taking a base $\{v_1, \ldots, v_n\}$ in V, we can consider A_1 and B_1 to be the matrices of linear transformations A and B of V. By a simple calculation it can be shown that the matrices A_1 and B_1, and so the transformations A and B, have characteristic polynomials $P(x)$ and $Q(x)$, respectively. Moreover, it is evident from (1.2) that rank $(A-B) = \text{rank } (A_1 - B_1) \leqslant 1$.

The above solution of problem Ia can be put in another useful form. Let V' be the vector space of those polynomials of $L[x]$ which have degree $\leqslant n-1$. By $v_i \to x^i (i=0, 1, \ldots, n-1)$ an isomorphism φ of V onto V' is induced. If under φ, A' and B' correspond to A and B, respectively, then

$$(1.3)\qquad A'x^i = x^{i+1} \quad (i=0, 1, 2, \ldots, n-2), \quad A'x^{n-1} = x^n - P(x)$$

and

$$(1.4)\qquad B'x^i = x^{i+1} \quad (i=0, 1, 2, \ldots, n-2), \quad B'x^{n-1} = x^n - Q(x).$$

In the sequel, we shall tacitly identify V and V' sometimes.

Next, we consider problem Ib. Let A, B be a solution of Ia, and let σ be an automorphism of V. Then $A^* = \sigma A \sigma^{-1}$, $B^* = \sigma B \sigma^{-1}$ is also a solution of Ia, as is readily verified. It seems reasonable, not to regard A, B and A^*, B^* as "different" solutions. Therefore, we shall say that Ia has a unique solution if the following is true: When A, B and A^*, B^* are solutions of Ia, then there exists an automorphism σ of V, such that $A^* = \sigma A \sigma^{-1}$ and $B^* = \sigma B \sigma^{-1}$.

Theorem 1.1. Problem Ia has a unique solution if and only if either $P(x)$ and $Q(x)$ are relatively prime or $P(x)=Q(x)$ is an irreducible polynomial.

Proof. Let either $P(x)$ and $Q(x)$ be relatively prime or $P(x)=Q(x)$ be an irreducible polynomial, and let A, B solve problem Ia. When H is the null-space of $A-B$, then

$$\Gamma=\{v|v \in H, Av \in H, \ldots, A^{n-2}v \in H\}$$

is a subspace of V and dim $\Gamma \geqslant 1$, for we have $\Gamma=\bigcap_{i=0}^{n-2}\{v|A^iv \in H\}$, and dim $\{v|A^iv \in H\} \geqslant n-1$. Let $v \neq 0$ be an element of Γ. As $Ax=Bx$ for all $x \in H$, we derive $Bv=Av$. Now $Av \in H$, and so $BAv=A^2v$, hence $B^2v=A^2v$. Repeating this argument, we find

$$Bv=Av,\ B^2v=A^2v,\ \ldots,\ B^{n-1}v=A^{n-1}v. \tag{1.5}$$

We shall prove that $v, Av, \ldots, A^{n-1}v$ are linearly independent. Assume that $v, Av, \ldots, A^{n-1}v$ are linearly dependent. Then a smallest $k (k=1, \ldots, n-1)$ exists such that A^kv is linearly dependent on $v, Av, \ldots, A^{k-1}v$. Therefore we can find $\lambda_0, \lambda_1, \ldots, \lambda_{k-1} \in L$, satisfying

$$A^kv=\lambda_0v+\lambda_1Av+\ldots+\lambda_{k-1}A^{k-1}v. \tag{1.6}$$

Introducing $p(x)=\lambda_0+\lambda_1x+\ldots+\lambda_{k-1}x^{k-1}-x^k$, we have $p(A)v=0$ from (1.6). Now we shall prove that $p(x)$ is a factor of $P(x)$. We can find polynomials $q(x)$ and $r(x)$ of $L[x]$, satisfying

$$P(x)=q(x)\,p(x)+r(x), \tag{1.7}$$

and either degree $r(x)<$ degree $p(x)$, or $r(x)=0$.

According to a theorem of Cayley we have $P(A)=0$. So (1.7) implies, as $p(A)v=0$,

$$r(A)v=P(A)v-q(A)\,p(A)v=0.$$

If $r(x) \neq 0$, we would have found a polynomial of degree less than k, satisfying $r(A)v=0$, contradicting our choice of $p(x)$. Therefore, $p(x)$ is a factor of $P(x)$. If $P(x)$ is irreducible, then $p(x)=P(x)$. Hence degree $p(x)=n$, a contradiction. If $P(x)$ and $Q(x)$ are relatively prime, we observe that $p(x)$ and $Q(x)$ are relatively prime. So we can find polynomials $a(x)$ and $b(x)$ in $L[x]$ satisfying

$$1=a(x)\,p(x)+b(x)\,Q(x). \tag{1.8}$$

Applying Cayley's theorem to the transformation B, we get $Q(B)=0$. So (1.8) implies $a(B)\,p(B)=I$ (the identity transformation). From (1.5) and (1.6) we derive that $p(B)v=0$, whence $v=a(B)\,p(B)v=0$. This contradicts our assumption $v \neq 0$, and it follows that $v, Av, \ldots, A^{n-1}v$ are linearly independent.

As $P(A)=Q(B)=0$, we have

$$A(A^{n-1}v)=A^nv-P(A)v,\ B(B^{n-1}v)=B^nv-Q(B)v. \tag{1.9}$$

Let φ be the isomorphism of V onto V' induced by $A^i v \rightarrow x^i (i=0, 1, \ldots, n-1)$' and let A' and B' correspond to A and B under φ. From (1.5) and (1.9) it is immediately clear that A' and B' satisfy (1.3) and (1.4). This proves the uniqueness of the solution of problem Ia.

In order to complete the proof of theorem 1.1 we now assume that $P(x)$ and $Q(x)$ have a non-constant common factor and that $P(x)$ and $Q(x)$ are not equal to the same irreducible polynomial. Then we can find non-constant polynomials $f(x)$, $p(x)$ and $q(x)$, such that $P(x)=p(x)\,f(x)$, $Q(x)=q(x)\,f(x)$, and such that $f(x)$ is relatively prime to at least one of $p(x)$, $q(x)$. Without any restriction we may assume that $p(x)$ and $f(x)$ are relatively prime. Moreover, we can assume that $f(x)$ is a monic polynomial, i.e. the coefficient of the highest power of x equals 1. We shall now construct two non-equivalent solutions of problem Ia. The linear transformations of both solutions are represented by matrices A, B and A, B^*. These matrices are defined by

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad B = \begin{pmatrix} A_1 & 0 \\ 0 & B_2 \end{pmatrix}, \quad B^* = \begin{pmatrix} A_1 & N \\ 0 & B_2 \end{pmatrix}.$$

A_1 is a matrix with characteristic polynomial $f(x)$. A_2 and B_2 are companion matrices of $p(x)$ and $q(x)$, respectively. However, if $p(x)=x+\alpha$ and $q(x)=x+\beta$, then A_2 and B_2 are 1×1 matrices, consisting of $-\alpha$ and $-\beta$, respectively. Finally, N is a matrix of which only the last column is different from 0. A, B and A, B^* are both solutions of problem Ia, as is readily verified. Suppose that A, B and A, B^* are equivalent. Then there exists a non-singular matrix X satisfying $XA=AX$, $XB=B^*X$. As A_1 and A_2 have relatively prime characteristic polynomials, we have

$$X = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}$$

where X_1 and X_2 are non-singular. (This can be proved as follows. Assume that

$$X = \begin{pmatrix} X_1 & X_3 \\ X_4 & X_2 \end{pmatrix}$$

Then from $XA=AX$ we derive $A_1X_3=X_3A_2$ and $A_2X_4=X_4A_1$. Since $f(x)$ and $p(x)$ are relatively prime, we can find polynomials $a(x)$ and $b(x)$ such that $1=a(x)f(x)+b(x)p(x)$. Substituting $x=A_2$ in the latter relation and noticing that $p(A_2)=0$ by Cayley's theorem, we find $a(A_2)f(A_2)=I$. Hence $X_4=a(A_2)f(A_2)X_4=a(A_2)X_4\,f(A_1)=0$, since $f(A_1)=0$ by Cayley's theorem. Similarly we can prove $X_3=0$). $XB=B^*X$ implies $NX_2=0$. As $N\neq 0$, and X_2 is non-singular, we have a contradiction. Hence, there exist two non-equivalent solutions of problem Ia. This completes the proof of theorem 1.1.

We shall also meet a multiplicative form of problem I. Once again, let V be an n-dimensional vector space over L, and let $f(x)$ and $g(x)$ be monic polynomials of degree n with coefficients in L. Finally, let $f(0)$ be $\neq 0$.

Problem II

a. To find: linear transformations A, B and C of V, related by $AC=B$ and satisfying the following conditions:
   (1) $f(x)$ and $g(x)$ are the characteristic polynomials of A and B, respectively.
   (2) There exists a subspace H of V with $\dim H \geqslant n-1$, such that the restriction of C to H is the identity transformation.

b. To find appropriate conditions for the uniqueness of the solution of II a.

There is a very simple connection between the problems I and II. Take $P(x)=f(x)$ and $Q(x)=g(x)$, and let A, B be a solution of problem Ia. Then A is non-singular, for $P(0)=f(0)\neq 0$. Define C by $C=A^{-1}B$. Then $C-I=A^{-1}(B-A)$. Therefore, rank $(C-I)\leqslant 1$. Hence, C satisfies condition 2 of problem IIa, and so a solution of this problem has been found. Conversely, let $P(x)$ $(P(0)\neq 0)$ and $Q(x)$ be given. Take $f(x)=P(x)$ and $g(x)=Q(x)$. If A, B and C form a solution of IIa, then A, B is a solution of Ia. Hence, there exists a one-to-one correspondence φ between the solutions of problem IIa and those of problem I in case that $P(0)\neq 0$. Using "different solutions" and "unique" for problem II as was done for problem I, we see that "being different" is preserved under φ. Hence, we have:

Theorem 1.2. Problem IIa has a solution which can be described as follows: When V is regarded as the linear space over L of polynomials $\in L[x]$ with degree $\leqslant n-1$, A and B are induced by

$$Ax^i=x^{i+1} \quad (i=0, 1, \ldots, n-2), \qquad Ax^{n-1}=x^n-f(x) \tag{1.10}$$

and

$$Bx^i=x^{i+1} \quad (i=0, 1, \ldots, n-2), \qquad Bx^{n-1}=x^n-g(x), \tag{1.11}$$

respectively, whereas $C=A^{-1}B$.

Problem IIa has a unique solution if and only if either $f(x)$ and $g(x)$ are relatively prime or $f(x)=g(x)$ is an irreducible polynomial.

It is clear that the relation between the problems I and II is trivial, and one might ask whether there is a point in introducing problem II at all. The answer is that the problens I and II appear in the analytic theory of hypergeometric functions in rather different situations which are not connected algebraically.

We shall deduce some consequences for future reference in the theory of hypergeometric functions. We make the special assumption that $f(x)$ ($f(0) \neq 0$) and $g(x)$ are relatively prime, and that both are products of linear polynomials $\in L[x]$:

$$f(x) = (x-\varrho_1)^{\mu_1} \dots (x-\varrho_r)^{\mu_r},\ g(x) = (x-\sigma_1)^{\nu_1} \dots (x-\sigma_s)^{\nu_s}, \tag{1.12}$$

where $\mu_1, \dots, \mu_r, \nu_1, \dots, \nu_s$ are positive integers satisfying

$$\mu_1 + \dots + \mu_r = \nu_1 + \dots + \nu_s = n. \tag{1.13}$$

$\varrho_1, \dots, \varrho_r, \sigma_1, \dots, \sigma_s$ are mutually different. We shall consider V to be the space of polynomials of $L[x]$ with degree $\leqslant n-1$. Then the linear transformations A, B and C, solving problem IIa, are described by theorem 1.2. Our purpose is to introduce new bases $\{a_1, \dots, a_n\}$ and $\{b_1, \dots, b_n\}$ in V, and to establish the relations between them.

If $p(x) \in V$, there are uniquely determined constants $\vartheta_1, \dots, \vartheta_n \in L$, satisfying

$$\frac{p(x)}{f(x)} = \frac{\vartheta_1}{x-\varrho_1} + \dots + \frac{\vartheta_{\mu_1}}{(x-\varrho_1)^{\mu_1}} + \frac{\vartheta_{\mu_1+1}}{x-\varrho_2} + \dots + \frac{\vartheta_n}{(x-\varrho_r)^{\mu_r}}. \tag{1.14}$$

Conversely, if $\vartheta_1, \dots, \vartheta_n$ are arbitrarily chosen in L, there exists a polynomial $p(x)$ of degree less than n, so that (1.14) holds. The correspondence $p(x) \leftrightarrow (\vartheta_1, \dots, \vartheta_n)$ is linear, hence $(\vartheta_1, \dots, \vartheta_n)$ can be regarded as coordinates in V. Let $\{a_1, \dots, a_n\}$ be the corresponding base of V. From (1.10), (1.11) and $C = A^{-1} B$ it is evident that all polynomials of degree $\leqslant n-2$ belong to H. As $H \neq V$ (otherwise $A = B$. This, however, cannot happen since $f(x)$ and $g(x)$ are relatively prime), H is the subspace of V of all polynomials with degree $\leqslant n-2$. We also derive from (1.10) and (1.11) that $Ap(x) = Bp(x) = xp(x)$ for all $p(x) \in H$, a fact which will be tacitly used in future.

In the next three theorems pairs of subscripts j, k and l, m will appear. Once and for all we agree that—apart from possible extra conditions—these subscripts take the values

$$\begin{cases} j = 1, \dots, r \quad ; \quad k = 1, 2, \dots, \mu_j, \\ l = 1, \dots, s \quad ; \quad m = 1, 2, \dots, \nu_l. \end{cases} \tag{1.15}$$

Furthermore, we shall use single subscripts i and h, each assuming the values $1, \dots, n$. A double subscript jk may be replaced by a single subscript i if and only if

$$i = \mu_1 + \dots + \mu_{j-1} + k, \text{ when } j > 1 \text{ or } i = k, \text{ when } j = 1. \tag{1.16}$$

Similarly, a double subscript lm may be replaced by h if and only if

$$h = \nu_1 + \dots + \nu_{l-1} + m, \text{ when } l > 1 \text{ or } h = m \text{ when } l = 1. \tag{1.17}$$

Now taking especially $p(x)=(x-\varrho_j)^{-k}f(x)=a_{jk}$ $(k\geqslant 2)$ in (1.14), we derive from the fact that $p(x)\in H$:

(1.18) $$a_{jk}\in H \qquad \text{if } k\geqslant 2.$$

In the same way we may deduce

(1.19) $$a_{j1}-a_{j'1}\in H.$$

We can also prove

(1.20) $$(A-\varrho_j I)a_{j1}=0,\ (A-\varrho_j I)a_{jk}=a_{j,k-1} \qquad (k\geqslant 2).$$

For take $p(x)=(x-\varrho_j)^{-1}f(x)$ in (1.14), then $p(x)=a_{j1}$. Using (1.10), and noticing that $\{(x-\varrho_j)^{-1}f(x)-x^{n-1}\}\in H$, we get

$$\begin{aligned}Aa_{j1}=Ap(x)&=A\{(x-\varrho_j)^{-1}f(x)-x^{n-1}\}+Ax^{n-1}=\\&=x\{(x-\varrho_j)^{-1}f(x)-x^{n-1}\}+x^n-f(x)=\\&=\varrho_j(x-\varrho_j)^{-1}f(x)=\varrho_j\,p(x)=\varrho_j a_{j1}.\end{aligned}$$

The second formula of (1.20) can be proved in a similar way.

Later on we shall need the following

Theorem 1.3. Let V be an n-dimensional vector space over L, H an $(n-1)$-dimensional subspace of V, $\{a_1, \ldots, a_n\}$ a base of V, and A a linear transformation of V satisfying (1.18), (1.19) and (1.20) (the ϱ's are mutually different). Then to every set $\{a_1^*, \ldots, a_n^*\}$ of elements of V which satisfies (1.18), (1.19) and (1.20) (if a_i is replaced by a_i^*) a constant λ can be found such that $a_1^*=\lambda a_1, \ldots, a_n^*=\lambda a_n$.

Proof. We first remark that this theorem has been enunciated without reference to problem II, being a general theorem on vector spaces and linear transformations. However, the notation suggests a relation to problem II. There is no j such that $a_{j1}\in H$. For assume that such a j exists. Then from (1.19) it would be clear that all $a_{j'1}\in H$, hence every $a_i\in H$. This contradicts the fact that $\{a_1, \ldots, a_n\}$ is a base of V and H an $(n-1)$-dimensional subspace of V. Let $\{a_1^*, \ldots, a_n^*\}$ be a set of vectors of V, satisfying (1.18), (1.19) and (1.20). As $a_{11}\notin H$, we can find $\lambda\in H$ satisfying $a_{11}^*-\lambda a_{11}\in H$. Let $\{a_1', \ldots, a_n'\}$ be defined by $a_{jk}'=a_{jk}^*-\lambda a_{jk}$. Then $\{a_1', \ldots, a_n'\}$ also satisfies (1.18), (1.19) and (1.20) (being homogeneous linear relations), whereas $a_{11}'\in H$. So we find $a_i'\in H$ for all values of i. To each value of j the eigenvectors of A corresponding to ϱ_j form a 1-dimensional subspace of V. [This is evident from the theory of Jordan matrices, but it can also be shown in the following way. Let a be a vector satisfying $(A-\varrho_1 I)a=0$. Then we can find constants $\lambda_{jk}\in L$ such that $a=\sum_{j,k}\lambda_{jk}a_{jk}$. Hence by (1.20)

$$0=(A-\varrho_1 I)\,a=\sum\lambda_{jk}(A-\varrho_1 I)\,a_{jk}=$$
$$=\sum_{k>1}\lambda_{jk}a_{j,k-1}+\sum\lambda_{jk}(\varrho_j-\varrho_1)\,a_{jk}=$$
$$=\sum_{k<\mu_j}\{\lambda_{jk}(\varrho_j-\varrho_1)+\lambda_{j,k+1}\}\,a_{jk}+\sum\lambda_{j\mu_j}(\varrho_j-\varrho_1)\,a_{j\mu_j}.$$

Now $\{a_1, \ldots, a_n\}$ is a base of V, so that we find

$$\lambda_{jk}(\varrho_j-\varrho_1)+\lambda_{j,k+1}=0 \qquad \text{if } k<\mu_j,$$
$$\lambda_{j\mu_j}(\varrho_j-\varrho_1)=0.$$

From these relations we see that $\lambda_{jk}=0$ when $j\neq 1$, and $\lambda_{1k}=0$ when $k>1$. Hence $a=\lambda_{11}a_{11}$.]

As a_{j1}' is an eigenvector of A corresponding to ϱ_j, we have $a_{j1}'=\lambda_j a_{j1}$ ($\lambda_j \in L$). However, a_{j1}' belongs to H, whereas $a_{j1} \notin H$ as shown above. So $\lambda_j=0$, and $a_{j1}'=0$ consequently.

Noticing that $(A-\varrho_j I)a_{j2}'=a_{j1}'=0$, we see that a_{j2}' is an eigenvector, whence it is a scalar multiple of a_{j1}. Hence, being a vector of H, a_{j2}' must be 0. Repeating these arguments, we can show that $a_{jk}'=0$ for all values of j and k. Therefore $a_{jk}^*=\lambda a_{jk}$ for all values of j and k. This completes the proof.

Now we introduce a base $\{b_1, \ldots, b_n\}$ of V, which is related to B as is $\{a_1, \ldots, a_n\}$ to A. If $p(x) \in V$ we can find $\xi_1, \ldots, \xi_n \in L$, so that

$$\text{(1.21)}\qquad \frac{p(x)}{g(x)}=\frac{\xi_1}{x-\sigma_1}+\ldots+\frac{\xi_{\nu_1}}{(x-\sigma_1)^{\nu_1}}+\frac{\xi_{\nu_1+1}}{x-\sigma_2}+\ldots+\frac{\xi_n}{(x-\sigma_s)^{\nu_s}}.$$

$\xi_1, \ldots, \xi_n$ can be considered as coordinates with respect to a base $\{b_1, \ldots, b_n\}$. Formulae analogous to (1.18), (1.19) and (1.20) can easily be derived:

$$\text{(1.22)}\qquad b_{lm} \in H \qquad \text{if } m \geqslant 2,$$

$$\text{(1.23)}\qquad b_{l1}-b_{l'1} \in H,$$

$$\text{(1.24)}\qquad (B-\sigma_l I)b_{l1}=0,\ (B-\sigma_l I)b_{lm}=b_{l,m-1} \qquad (m\geqslant 2).$$

The relations connecting $\{a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$ will now be established.

Take $p(x)=b_{lm}=(x-\sigma_l)^{-m}g(x)$ in formula (1.14). Then the coefficient $\vartheta_{jk,lm}$ of $(x-\varrho_j)^{-k}$ in (1.14) can easily be calculated:

$$\text{(1.25)}\qquad \vartheta_{ih}=\vartheta_{jk,lm}=\frac{1}{(\mu_j-k)!}\left[\frac{d^{\mu_j-k}}{dx^{\mu_j-k}}\frac{g(x)\,(x-\varrho_j)^{\mu_j}}{f(x)\,(x-\sigma_l)^m}\right]_{x=\varrho_j}.$$

In a similar way, taking $p(x)=a_{jk}=(x-\varrho_j)^{-k}f(x)$ in (1.21), we find for the coefficient $\xi_{lm,jk}$ of $(x-\sigma_l)^{-m}$ in (1.21)

$$\text{(1.26)}\qquad \xi_{hi}=\xi_{lm,jk}=\frac{1}{(\nu_l-m)!}\left[\frac{d^{\nu_l-m}}{dx^{\nu_l-m}}\frac{f(x)\,(x-\sigma_l)^{\nu_l}}{g(x)\,(x-\varrho_j)^k}\right]_{x=\sigma_l}.$$

We summarize these results in

Theorem 1.4. The bases $\{a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$ of V are connected by

$$\{b_1, \ldots, b_n\} = \{a_1, \ldots, a_n\}\,\Theta \tag{1.27}$$

$$\{a_1, \ldots, a_n\} = \{b_1, \ldots, b_n\}\,\Xi \tag{1.28}$$

where the matrices $\Theta = (\vartheta_{ih})$ and $\Xi = (\xi_{hi})$ are given by (1.25) and (1.26), respectively.

Finally, we shall pay some attention to the 1-dimensional subspace $(C-I)V$. Let c generate this subspace. Then $(c) = (C-I)V = (A^{-1}B-I)V'$. From (1.10) and (1.11) it is evident that $(A^{-1}B-I)x^i = 0$ when $i = 0, \ldots, n-2$, and $(A^{-1}B-I)x^{n-1} = A^{-1}(x^n - g(x)) - x^{n-1}$. Let ϱ and σ be defined by

$$\varrho = (-1)^n f(0) = \varrho_1^{\mu_1} \ldots \varrho_r^{\mu_r}, \quad \sigma = (-1)^n g(0) = \sigma_1^{\nu_1} \ldots \sigma_s^{\nu_s}. \tag{1.29}$$

It is easily seen that $A^{-1}1 = x^{-1}(1 + (-1)^{n-1}\varrho^{-1}f(x))$, and

$$A^{-1}\varphi(x) = x^{-1}(\varphi(x) - \varphi(0)) + \varphi(0)A^{-1}1 \quad \text{for every } \varphi(x) \in V.$$

Hence,

$$\begin{aligned}(A^{-1}B-I)x^{n-1} &= A^{-1}(x^n - g(x)) - x^{n-1} = x^{-1}(x^n - g(x) + g(0)) - \\ &\quad - g(0)A^{-1}1 - x^{n-1} = -x^{-1}g(x) + x^{-1}g(0) - x^{-1}g(0)(-1)^{n-1}\varrho^{-1}f(x) = \\ &= x^{-1}(\varrho^{-1}\sigma f(x) - g(x)).\end{aligned}$$

So we can take

$$c = x^{-1}(\varrho^{-1}f(x) - \sigma^{-1}g(x)). \tag{1.30}$$

Finally, we determine the coordinates of c with respect to the bases $\{a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$.

Putting $p(x) = c = x^{-1}(\varrho^{-1}f(x) - \sigma^{-1}g(x))$ in (1.14), we find for the coefficient δ_{jk} of $(x-\varrho_j)^{-k}$

$$\delta_{jk} = \frac{1}{(\mu_j - k)!}\left[\frac{d^{\mu_j-k}}{dx^{\mu_j-k}}\,\frac{(\varrho^{-1}f(x) - \sigma^{-1}g(x))\,(x-\varrho_j)^{\mu_j}}{xf(x)}\right]_{x=\varrho},$$

whence

$$\delta_{jk} = \frac{-1}{(\mu_j - k)!}\left[\frac{d^{\mu_j-k}}{dx^{\mu_j-k}}\,\frac{g(x)\,(x-\varrho_j)^{\mu_j}}{\sigma f(x)\,x}\right]_{x=\varrho_j}. \tag{1.31}$$

In a similar way the coefficient ε_{lm} of $(x-\sigma_l)^{-m}$ in (1.21) can be calculated, when $p(x) = c$ is substituted. The result is

$$\varepsilon_{lm} = \frac{1}{(\sigma_l - m)!}\left[\frac{d^{\sigma_l-m}}{dx^{\sigma_l-m}}\,\frac{f(x)\,(x-\sigma_l)^{\nu_l}}{\varrho g(x)\,x}\right]_{x=\sigma_l}. \tag{1.32}$$

Theorem 1.5. The subspace $(C-I)V$ is generated by a vector c, and

$$c = \delta_1 a_1 + \ldots + \delta_n a_n = \varepsilon_1 b_1 + \ldots + \varepsilon_n b_n, \tag{1.33}$$

where $\delta_1, \ldots, \delta_n$ and $\varepsilon_1, \ldots, \varepsilon_n$ are given by (1.31) and (1.32), respectively.

Moreover, the following formula holds

$$(1.34) \qquad \delta_{11}+\delta_{21}+\ldots+\delta_{r1}=\varepsilon_{11}+\varepsilon_{21}+\ldots+\varepsilon_{s1}=\varrho^{-1}-\sigma^{-1}.$$

The first part of the theorem follows from the foregoing exposition. In formula (1.34) $\delta_{11}+\ldots+\delta_{r1}$ can be interpreted as the sum of the residues of $x^{-1}(\varrho^{-1}f(x)-\sigma^{-1}g(x))f^{-1}(x)$. This sum is easily evaluated.

REFERENCES

1. THOMAE, J., Über die höheren hypergeometrischen Reihen. Math. Ann., 2, 427–444 (1870).
2. NØRLUND, N. E., Hypergeometric Functions. Acta Math., 94, 289–349 (1955).
3. BIRKHOFF, G. and S. MAC LANE, A survey of modern algebra. The MacMillan Company, New York.

(To be continued)